Assumptions and Rules:

1. The ghosts come out from the center box into board through the only exit. Then, pick one direction randomly to start.
2. The ghosts only change its direction when it hit a wall.
3. When the ghost hit a wall, it changes its direction randomly.
4. Pacman start from a certain position and the moving direction is left.
5. Pacman keep moving until it hit a wall.
6. Pacman change its direction if agent assign a legal action
7. There is not fruit, and number of live of Pacman
8. In general mode, which Pacman did not eat power pellet, ghosts can kill the Pacman when they catch the Pacman, and the speed of the ghosts are as same as the Pacman. In reversed mode, which Pacman eat power pellet, Pacman can kill the ghosts, and the speed of the ghosts is slower. This state is limited by certain time.
9. After the ghost being killed, it disappears, and appears in center box immediately, after certain time, it will come out of box.
10. When there is not dot and power pellet in the board, agent win. When Pacman is killed by the ghosts, agent lost.

State Space(S):

State space is a set of all possible states. For each state, it consists of cells of a grid, which is 'G', and time, which is 'T', including several elements:

1. Empty cell: It is represented by value 0, which means there is nothing in this cell. The ghosts and the Pacman can move into it.
2. Wall: It is represented by value -1, which means there is wall in this cell. The ghosts and the Pacman cannot move into it, and also when they hit it, they change direction following the rules.
3. Dot: It is represented by value 1, which means there is dot in this cell.
4. Power pellet: It is represented by value 2, which means there is power pellet in this cell.
5. Pacman: Position of the Pacman, which is P_p; the direction of the Pacman, which is D_p; the mode of the Pacman, which is M_p; the timer of reversed mode which is C_p.
6. Ghosts: Position of the ghost, which is P_{gi} where i is the number of ghost; the direction of the ghost, which is D_{gi}; the mode of the ghost, which is M_{gi}; the timer of reversed mode, which is C_{gi}. For example, the state of the first ghost is represented by $\{P_{g1}, D_{g1}, M_{g1}, C_{g1}\}$

Action Space(A):

There are four actions of agent: Up, Down, Left, Right. (Although the Pacman can stop moving, it still have direction. Because the direction of it is to a wall, it cannot move. Once the agent assigns a legal action for it, it will move to that direction.)

Transition Function(T):

Transition function is compute the current state S^t to new state S^{t+1}
Dot: In current state S^t, if the Pacman move into the cell of the dot, it eat the dot, and in the new

state S^{t+1}, the value of this cell update to 0, which is empty.

Power Pellet: In the current state S^t, if the Pacman move into the cell of the power pellet, it eat the power pellet, and in the new state S^{t+1}, the value of this cell update to 0. At the same time, the mode of the Pacman and the ghost, which is M_p and M_{gi}, update to reversed mode, M_p = r and M_{gi} = r, and the reversed mode timer start from 25-time steps, which is C_p and C_{gi}, C_p = 25, C_{gi} = 25.

(M = [g, r], where g is general mode, r is reversed mode. C = [0, 25], C is decreased 1 by each time step unless it is 0)

Ghost: In the current state S^t,

if D_{gi} is left, then V' = $V_{r-1,c}$;

if D_{gi} is right, then V' = $V_{r+1,c}$;

if D_{gi} is up, then V' = $V_{r,c-1}$;

if D_{gi} is down, then V' = $V_{r,c+1}$

V' is the value of the new cell which the ghost will move into in new state S^{t+1}

If the V' is not -1, then the ghost moves into the new cell in new state S^{t+1}

if the V' is -1, which means there is wall, then the ghost chooses a new direction randomly D', D = D'.

In new state S^{t+1}, if there is Pacman in the same cell, and the M_{gi} is g, the mode is general mode then the game over; if there is Pacman in the same cell, and the M_{gi} is r, the mode is reversed mode, then this ghost is killed and disappear, and appear immediately in the center box.

Pacman: In the current state S^t,

if D_p is left, then V' = $V_{r-1,c}$;

if D_p is right, then V' = $V_{r+1,c}$;

if D_p is up, then V' = $V_{r,c-1}$;

if D_p is down, then V' = $V_{r,c+1}$

V' is the value of the new cell which the Pacman will move into in new state S^{t+1}

If the V' is not -1, then the Pacman moves into the new cell in new state S^{t+1}

if the V' is -1, which means there is wall, then the Pacman stop moving at that cell unless the legal direction has been assigned.

In new state S^{t+1}, if there is a ghost in the same cell, and the M_p is g, the mode is general mode then game over; if there is a ghost in the same cell, and the M_p is r, the mode is reversed mode, then the Pacman eat the ghost at that cell.

The timer of reversed mode C:

If the Pacman eat a power pellet, then C_p = 25, C_{gi} = 25, and mode M_p = r, M_{gi} = r in the new state S^{t+1}. Then, C_p and C_{gi} is decreased 1 in every new state S^{t+1} until C is 0. If C is not 0, mode M_p and M_{gi} always are reversed mode. If C is 0, the mode M_p and M_{gi} are updated to general mode.

Reward Function(R):

There are three ways to get reward:

If the Pacman eat a dot or power pellet, then R(S) = 1;

If the Pacman eat a ghost, then R(S) = 100;

If there are not dots and power pellet on the board, then win the game, and R(S) = 1000;

If not the conditions above, R(S) = 0;

Score = Score + R(S).